

Comprehensive Systematic Search Process of Health Literature: Hunting Pearls out of the Sea

Ahmed S¹, Vaska M², Turin TC^{1,3}

¹Department of Family Medicine, University of Calgary, Calgary, Alberta, Canada. ²Knowledge Resource Service, Alberta Health Services, Calgary, Alberta, Canada. ³Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada.

Abstract:

Background: Clinical information is expanding at an accelerated rate all over the world, encompassing scholarly articles, books, practice guidelines and grey literature. In knowledge synthesis practices systematic reviews are seen as the gold standard and in the heart of this methodology is a “systematic search”.

Methods: This paper focuses on describing the “systematic search” process, which is different than a general-purpose search in conventional search engines. Our previously published paper was used as the example while describing the search process health care researchers might follow.

Results: Almost every systematic search and screening reported in peer-reviewed scientific journals involves the following major steps: Generate appropriate and comprehensive sets of search words, find appropriate databases for the search, search in databases with a combination of search words, export search results into reference management software, and screening of literature.

Conclusion: In this paper, we tried to focus on the systematic searching and literature screening for knowledge synthesis focusing on medical or health research. All of the methodological steps described here are part of standard practice for synthesis among scientific communities, while also based on the experiences of the authors in conducting review studies.

(JNHFB 2016; 5 : 12-16)

Introduction

Clinical information is expanding at an accelerated rate all over the world, encompassing scholarly articles, books, practice guidelines and grey literature in the form of academic dissertations, magazines, blogs and Internet sites. Utilization of this information into practice is a prime concern among researchers and practitioners¹. Nonetheless it is not possible for anyone, even for an expert in any discipline to know all of the latest information in that discipline. In addition, the reliability and evidence-into-practice of any information is always questionable, especially in the medical field². To mitigate these issues, knowledge synthesis practice has been considered at length, with systematic reviews seen as the gold standard in the knowledge synthesis arena³. Over time, other types of review articles have been developed, such as scoping reviews, realist reviews, qualitative systematic reviews, etc. where each has its unique features and acceptability

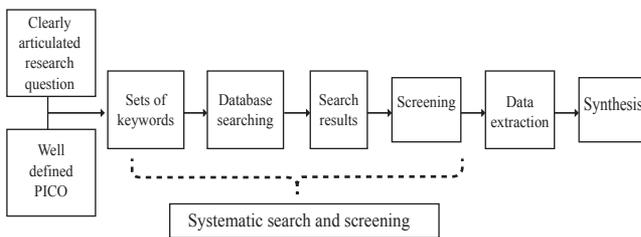
into knowledge synthesis⁴. The fundamental reason for the prioritizing and overwhelming acceptance of synthesized knowledge in the form of systematic, scoping, or realist reviews is their meticulous methodology; the heart of this methodology is a “systematic search”.

A “systematic search” is different than a general-purpose search in conventional search engines (Google, Yahoo, Bing etc.). Undoubtedly, conventional search engines will produce hundreds, even thousands of results for any word or phrase. However, specificity and stability of the conventional search results over time are questionable⁵. On the other hand, a systematic search consists of specific sets of search words according to the research question, searching in specialized databases designed for scientific literature⁶. These specialized databases have higher precision, superior stability, integrated export functions, a user account service for saving search history, and a better interface to express complex search strings⁷. It should be noted that a systematic search will generate results, but to take these results into the synthesis phase, rigorous screening of the results is also important. Figure 1 shows the positioning of a systematic search and screening in the synthesis process.

Corresponding Author

Tanvir Chowdhury Turin
Department of Family Medicine, Room G012F, Health Sciences Center,
3330 Hospital Drive Northwest, Calgary, Alberta T2N 4N1, Canada.
Email: turin.chowdhury@ucalgary.ca

Figure 1: Overview of systematic synthesis of literatures highlighting systematic search and screening.



Steps of systematic searching and screening:

Almost every systematic search and screening reported in peer-reviewed scientific journals involves the following major steps:

1. Generate appropriate and comprehensive sets of search words
2. Find appropriate databases for search
3. Search in databases with a combination of search words
4. Export search results into reference management software
5. Screening of literature

These steps are described below with the example of a recently published review from our research group. This review addressed the question of barriers to accessing primary healthcare for Canadian immigrants⁸.

Generate appropriate and comprehensive sets of search words:

Search words are usually two types: (i) subject headings and (ii) keywords. For the purpose of medical literature searching, subject headings are known as MeSH (Medical Subject Heading) terms. MeSH terms are organized terminology to indexing and cataloging biomedical and health related information in the National Library of Medicine catalog and databases⁹. Specific MeSH terms are assigned by the indexers to describe the content of articles. Keywords, on the other hand, are also known as “text-words” or “free text words”. Keywords are basically the words or phrases that researchers choose to search a database. Keyword searching is required when a researcher is interested in any new topic for which MeSH terms have not yet been assigned or to search for a very specific aspect such as a country name¹⁰.

It is always wise to break down a well-defined research question into parts and look for search words for each of the parts. For example, in the above mentioned review article, the major parts are barriers, primary healthcare, immigrants and Canada. Then, during a brainstorming session, as many synonyms as possible for each of the search words are identified. All of these synonymous words are searched in the MeSH database of NLM and EMBASE to identify whether these are MeSH terms or not. When a word or phrase is recognized as a MeSH term, the scope of that word or phrase can be identified from the database. In addition,

the search words are usually reported in the actual synthesis as keywords or MeSH terms. Sometimes words or truncated form of words are written with an asterisk (*) after it. This means the search engine will include any terms that start with the word preceded by the asterisk¹¹. This is like adding a wild card to the search term. Truncation is also used if any word has multiple spelling variations. Comprehensive search word sets of the review are shown in Table -1.

Table 1: Search keywords in details used in the example paper

Keywords for barrier:

barrier* [Keyword]; factor* [Keyword]; risk* [Keyword]; risk [MeSH]; “risk factor*” [Keyword]; risk factors [MeSH]; Prejudice [Keyword, MeSH]; self-conscience* [Keyword]; issue* [Keyword]; attitude* [Keyword]; attitude [MeSH]; uncertainty [Keyword, MeSH]; mistrust [Keyword]; obstacle* [Keyword]; hurdle* [Keyword]; difficulty [Keyword]; obstruction [Keyword]; impediment [Keyword]; Challenge* [Keyword]; confront* [Keyword]; defy [Keyword]; defiance [Keyword]; object* [Keyword]; contest* [Keyword]; oppos* [Keyword]; question* [Keyword]

Keywords for primary healthcare and family physician:

“Primary Healthcare” [Keyword]; delivery of primary healthcare [Keyword, MeSH]; “primary health care” [Keyword]; “primary healthcare system*” [Keyword]; “primary health care system*” [Keyword]; “public health” [Keyword, MeSH]; “primary healthcare seeking behavior” [Keyword]; “primary health care seeking behaviour” [Keyword]; “primary health behaviour” [Keyword, MeSH]; “family physician” [Keyword, MeSH]; “family doctor” [Keyword]; “family medicine” [Keyword]

Keywords for immigrants:

Immigrant* [Keyword]; emigrant* [Keyword]; alien* [Keyword]; emigrants and immigrants [MeSH]; Newcomer [Keyword]

Adapted with permission from Ahmed S, Shommu NS, Rumana N, Barron GR, Wicklum S, Turin TC. Barriers to access of primary health care by immigrant populations in Canada: A literature review. J Immigr Minor Health 2015 Sep 12. [Epub ahead of print]

Find appropriate database for search:

There are numerous specialized databases for searching the literature. However, each of these databases has its own focus and principle for indexing literature. For instance, MEDLINE provides extensive coverage of the international literature on biomedicine, including the allied health fields and the biological and physical sciences, humanities, and information science as they relate to medicine and health care¹². In contrast PsycINFO is a specific database that covers the professional and academic literature in psychology and related disciplines¹³.

Therefore, it is crucial to identify the appropriate databases in which the indexing strategy matches the research question. It is thus recommended to search as many databases as feasible to adequately address a research question,

despite the possibility that that this will generate potential duplication in the search results returned. Aside from a peer-reviewed literature search, a search of the grey literature is becoming a common trend in current systematic syntheses. Grey literature may be considered nonconventional, and is often associated with rapid publication. Grey documents may include: preprint reports, preliminary progress reports, advanced reports, technical reports, statistical reports, memos, market research reports, theses, conference proceedings, technical specifications and standards, non-commercial translations, bibliographies, technical and commercial documentation, and government reports and documents¹⁴. Similar to published literature, there are specific sources for a grey literature search. However, it should be mentioned that grey literature resources are still in their infancy; moreover, their precision and searching mechanism are not as developed as the specialized databases for peer-reviewed literature¹⁵. As an example of a standard way of reporting the databases in actual synthesis, Table-2 has been taken from the above-mentioned review article.

Table-2: List of databases searched to identify literature for this synthesis.

For published article:	
<ul style="list-style-type: none"> • MEDLINE • EBM Reviews (including Cochrane) • PubMed • Scopus • CINAHL Complete • Family & Society Studies Worldwide • Health Source - Consumer Edition • Psychology and Behavioral Sciences Collection • SocINDEX with Full Text 	<ul style="list-style-type: none"> • EMBASE • PsycINFO • PubMed Central • Web of Science • Academic Search • Family Studies Abstracts • Health Source: Nursing/Academic Edition • Social Work Abstracts
For grey literature:	
<ul style="list-style-type: none"> • Google • Google Scholar • ProQuest (theses & dissertations) • OpenDOAR (institutional repositories) • Health Sciences Online (HSO) • Turning Research into Practice (TRIP) 	<ul style="list-style-type: none"> • OAIster (World Cat) • Canadian Institute for Health Information • Public Health Agency of Canada (PHAC) • Health Canada • National Institutes of Health (NIH)

Adapted with permission from Ahmed S, Shommu NS, Rumana N, Barron GR, Wicklum S, Turin TC. Barriers to access of primary health care by immigrant populations in Canada: A literature review. J Immigr Minor Health 2015 Sep 12. [Epub ahead of print]

Search in databases with combination of search words:

Combinations of search words are important when searching any database. The reason for a combination is to make sure that none of the keywords or MeSH terms are missing from the search. Keywords are usually combined with Boolean operators (OR, AND, NOT)¹⁶. Words or sets of words combined with “OR” will return results where any of the words are present, thus expanding the search yield. Combinations with “AND” will return results where representation of all sets or words are present, thereby narrowing the search return yield. “NOT” is used to omit any word or sets of words in a search. The use of a Boolean operator can be shown according to Figure-2. For the example of the review mentioned before, all search words for each set (“barriers”, “primary healthcare”, “immigrants” and “Canada”) were joined with “OR”. Then individual sets were joined together by “AND”. Therefore, the search results should comprise literature that contains at least a word from the “barrier” set, a word from the “primary healthcare” set, a word from the “immigrants” set and the word “Canada”. Each database has its own search strategy. For some databases, the researcher can input the entire search string joined by a Boolean operator, while in other databases, each of the search words are represented individually, and joined together to produce the final search string(s).

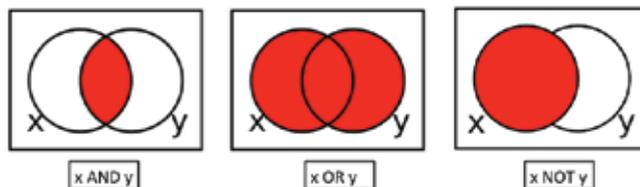


Figure 2: Representation of Boolean operators: AND, OR, NOT. x and y represents two sets where red is showing the output of combining x and y with the operators.(This figure is taken from Google Image labeled for reuse)

Export search results into reference management software:

Since multiple database searching is quite common, there is a need to store all the results in a single place before screening. Reference management (RM) software such as RefWorks, Endnote, Papers, CiteUlike, Zotero, WizFolio etc. serve this purpose. RM software is also known as bibliographic software, citation management software, or personal bibliographic file managers. Nearly every database has a direct export option for transferring search results into any desired RM software program. Citations and abstracts of the articles are exported into the RM software directly from the searched databases. RM software comes with multiple options 10 such as organizing

references by creating folders, duplicate identification and removal, compatibility with word processors, searching the full text of a PDF, a data exchange capability with other RM software, formatting citations in multiple styles, as well as exporting those citations into multiple formats such as tab delimited or XML.

Screening of literature

Screening of literature involves the following steps:

- Removal of duplicates/multiple copies of identical literature
- Developing PICO and inclusion criteria
- Screening based on title and abstract
- Screening based on full-text

These steps are briefly explained below:

Removal of duplicates/multiple copies of identical literature

Due to searching in multiple databases, several copies of the same citation may accumulate in the RM software. Screening out and removal of duplicates / additional copies is the first step in screening process. Most of the RM software such as RefWorks, Endnote, Papers etc has built-in duplicate identification and removal functions. However, it should be noted that the algorithms that RM software use for duplicate removal is not one hundred percent efficient.

Therefore, duplicates that are not filtered out through RM software functions need to be removed manually. Duplicate-free references can be exported to any spreadsheet, which makes further screening easier.

Developing PICO and inclusion criteria:

PICO is the acronym of population, intervention, comparison and outcome. PICO is effective for systematic reviews of intervention studies, hence quite common for quantitative synthesis. Depending on the study question and focus of the review, variants of PICO are used to put additional limits on the search results and study selection¹⁷. For instance, PICOS (where S stands for study design) or PICOT (where T stands for time frame) are two variants of PICO. For quantitative reviews, PICO is not only important to articulate the study question and search strategy, but is also essential in developing inclusion criteria for study selection. Inclusion criteria for literature screening towards a synthesis will be based on the correct population, appropriate intervention, valid comparison, and desired outcome. It should be noted that a systematic search is also the basis of qualitative and narrative reviews. In most the qualitative studies and some mixed-method studies, comparison is out of scope and clearly PICO does not completely fit into these scenarios¹⁸. Qualitative studies inherently are not destined to compare the control and intervention groups, but to elaborate on answers for health service research questions. Unfortunately, there is no highly tested and appropriate tool like PICO for qualitative synthesis. Cooke et al proposed an alternative

tool for qualitative synthesis called SPIDER (S=sample, PI=phenomenon of interest, D=Design, E=Evaluation, R=Research type) but it has yet to be tested on a wider range of topics¹⁹. The review mentioned here is a qualitative synthesis. In that review, the population is immigrants of Canada, and the outcome pertains to barriers accessing primary healthcare. There is no intervention and comparison. Based on this partial PICO, that review considered the following inclusion criteria for literatures:

- The study was conducted on Canadian immigrants.
- The study focused on the barriers to accessing primary healthcare in a Canadian context.

Screening based on title and abstract:

Sets of non-duplicated literature are usually screened in two steps. The first step is screening based on title and abstract. Reviewers typically go through the titles and abstracts of the pulled literature and look for the inclusion criteria in the abstract. If an abstract meets the inclusion criteria, that article gets selected for a full-text read. Screening through titles and abstracts also reduces the volume of full text reads.

Screening based on full text reads:

Full texts of the articles selected by title and abstract screening are downloaded and read by reviewers. As per the previous strategy, reviewers try to identify the inclusion criteria in the full texts of the articles. Articles that meet the criteria are taken for data extraction and synthesis. Data are usually extracted in tabular format along with other information such as author name, date of publication, study type, etc. Extracted data are commonly described into the synthesis.

Where appropriate, reviewers also assess the quality of the final selected studies, which increases the acceptability of the synthesis among scientific communities. Figure-3 shows the overall flow of systematic literature searching and screening as described in this literature.

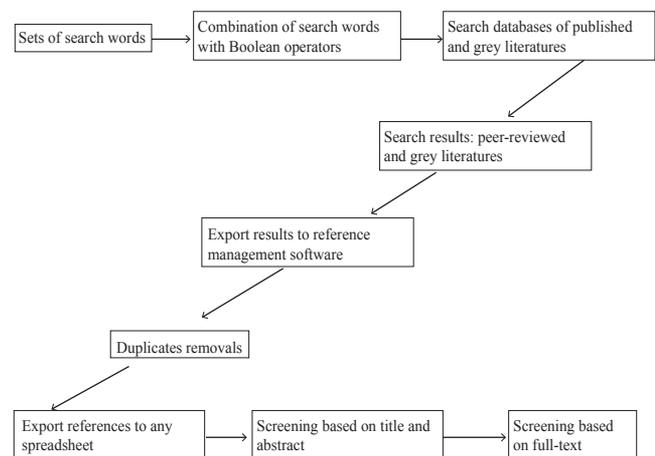


Figure 3: Flowchart representing the key steps in systematic search for literature.

Summary

To achieve the highest level of competency in any field, possession of relevant and up-to-date information is of utmost importance. Synthesis of information is necessary when embarking on any research project. Well-organized information is thus fundamental for evidence-based practice. Therefore, systematic searching and screening of literature is common in clinical practice and health research. People who want to keep up with new information in their field must practice systematic searching and screening as part of their professional life. Although the context of any project or one's interest may vary, the strategy of systematic searching and screening of information will remain similar. The fundamentals always lay in appropriate search words, correct selection of databases, and a rigorous selection method based on inclusion criteria structured around PICO.

Closing remarks

In this paper, we tried to focus on the systematic searching and literature screening for knowledge synthesis focusing on medical or health research. All of the methodological steps described here are part of standard practice for synthesis among scientific communities, while also based on the experiences of the authors in conducting review studies. This article is intended for beginning healthcare professionals, students and researchers who want to synthesize the wealth of scientific information available in their research and practice through a systematic way.

Considerations for practice

- Try to make the list of search words as comprehensive as possible
- Search all appropriate databases but consider the volume of overlapping / duplication with increased number of databases.
- Carefully use Boolean operators to combine search words.
- Traditional PICO tool may not fit entirely for all types of syntheses.

References

1. Haynes B and Haines A. 10 Barriers and bridges to evidence-based clinical practice. *Getting Research Findings into Practice* 2008; 115.
2. Rosenberg W and Donald A. Evidence based medicine: an approach to clinical problemsolving. *BMJ: British Medical Journal* 1995; 310, p. 1122.
3. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Annals of internal medicine* 1997; 126: 376-380.
4. Grant MJ and Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal* 2009; 26: 91-108.
5. Boeker M, Vach W, Motschall E. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC medical research methodology* 2013; 13: 131.
6. Tranfield DR, Denyer D, Smart P. Towards a methodology for developing evidenceinformed management knowledge by means of systematic review. *British journal of management* 2003; 14: 207-222.
7. McGowan J and Sampson M. Systematic reviews need systematic searchers. *Journal of the Medical Library Association* 2005; 93: 74.
8. Ahmed S, Shommu NS, Rumana N, Barron GR, Wicklum S and Turin TC. Barriers to Access of Primary Healthcare by Immigrant Populations in Canada: A Literature Review. *Journal of Immigrant and Minority Health* 2015; 1-19.
9. U. S. N. L. O. Medicine. Medical Subject Headings. Available: <https://http://www.nlm.nih.gov/mesh/>
10. Thompson J, Davis J, Mazerolle L. A systematic method for search term selection in systematic reviews. *Research Synthesis Methods* 2014; 5: 87-97.
11. Bartels EM. How to perform a systematic search. *Best Practice & Research Clinical Rheumatology* 2013; 27: 295-306.
12. Fact Sheet: MEDLINE. Available: <https://http://www.nlm.nih.gov/pubs/factsheets/medline.html> (15 Nov 2015).
13. PsycINFO.
14. Alberani V, Pietrangeli PDC, Mazza A. The use of grey literature in health sciences: a preliminary survey. *Bulletin of the Medical Library Association* 1990; 78: 358.
15. Rothstein H. R and Hopewell S, *Grey literature* 2009.
16. SCOPUS. Boolean Operators. Available: http://help.scopus.com/Content/h_bscsrch_boolop.htm (15 Nov 2015).
17. Methley AM Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC health services research* 2014; 14: 579.
18. Jones K. Mission drift in qualitative research, or moving toward a systematic review of qualitative studies, moving back to a more systematic narrative review. *Qualitative Report* 2004; 9: 95-112.
19. Cooke A, Smith D, Booth A. Beyond PICO The SPIDER Tool for Qualitative Evidence Synthesis. *Qualitative health research* 2012; 22: 1435-1443.